

# Är det möjligt att utvärdera kodkvalitet med hjälp av machine learning?

POPULÄRVETENSKAPLIG SAMMANFATTNING ARVID MALMSTRÖM & ERIK KINDT

I en framtid där machine learning samt artificiell intelligens kommer att ta mer och mer plats så finns det alltid en fråga som kvarstår. I vilka användningsområden finns det ett behov av machine learning och var går det att tillämpas? Berört arbete har undersökt huruvida machine learning kan användas för att utvärdera kvalitet av skriven källkod.

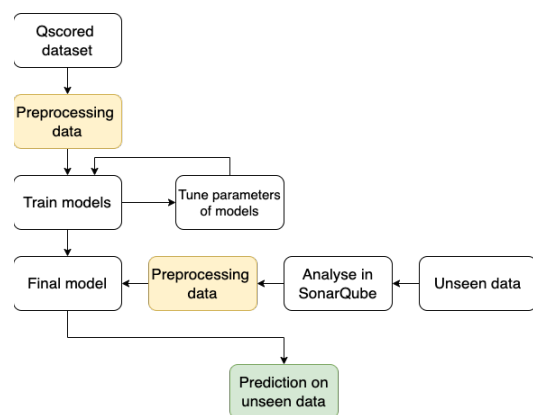
För stora företag som sysslar med mjukvaru-utveckling så är en viktig aspekt att utvecklarna skriver högkvalitativ källkod. Detta leder till att det är lätt att testa programmet, lätt att förstå programmet samt att det blir lätt att göra behövda förändringar i programmet. Frågan är hur man ska mäta en sådan kodkvalitet.

Arbetet ifråga angriper detta problem genom machine learning(ML). Examensarbetet har undersökt om det är möjligt att bygga ML-modeller för att mäta hur bra skriven koden är, både för enskilda filer men också för hela projekt som inkluderar flertalet filer.

Det byggdes flera modeller, delvis en som ska vara bättre på små, enskilda filer och delvis en som ska vara bättre på att modellera stora projekt. I modellerna i fråga så fanns det fem in-parametrar, vilka var: Antal rader, antal metoder, antal 'code smells', procentsats av hur mycket duplicerad kod det fanns samt cyklomatisk komplexitet. Dessa parametrar användes för att bygga fler stycken modeller för att prova flera olika ML-algoritmer.

De algoritmer som producerade bäst resultat var Random forest samt Artificial neural networks. För att sedan prova dessa modeller i verkligheten så konstruerades två tester.

Hela arbetsprocessen visas i figuren nedan.



Test 1 jämförde modellens satta betyg på nya filer med erfarna mjukvaru-utvecklare på ett stort svenskt företag. Resultaten från test 1 påvisade att både Random forest och ANN gav de nya filerna samma resultat i stor utsträckning. Modellen höll med utvecklarna mer när det kom till front end filer än med back end filer.

Test 2 påvisade att det är möjligt att använda modellerna för att utvärdera ett stort projekt under en lång tid för att se om kodkvaliteten blev bättre.

Slutsatserna som dras efter arbetet är att det finns potential för att använda ML för att utvärdera kodkvalitet. Det finns vissa svårigheter att hitta passande data, men skulle sådana data set konstrueras så är det absolut en väg framåt.